Indian J. Fish., ***49***(4) : 361-368, Oct.-Dec., 2002

Evaluation of Schaefer's production model

M. SRINATH

Central Marine Fisheries Research Institute, Cochin - 682 014, India

ABSTRACT

Schaefer's production model is one of the most widely used surplus production models employed in fish stock assessment. This model leads to estimation of the management reference points, namely, the Maximum Sustainable Yield (MSY) and the corresponding optimal fishing effort (fMSY). In this paper, the performance of the process error estimators of the parameters of the two non-equilibrium formulations of the Schaefer's production model in the presence of different magnitudes of the observation and process errors are investigated. The choice of the appropriate formulation vis-a-vis the estimates of the parameters and the management reference points is also discussed.

Introduction

In fish stock assessment, the dynamics of the exploited fish populations is studied by application of micro and macro production models. The micro models (the age structured or size structured models) take into account the individual parameters that govern the dynamics of the stock, namely, growth, recruitment, mortality (natural and fishing) and age at first capture. Thus, these are data intensive and require a lot of inputs. From this, models are then formulated to relate the yield from the stock as a function of these parameters taken either singly or in combination with each other. From such a functional relationship appropriate management options are estimated. Macro production models (such as Schaefer surplus production model) deal only with interrelationship of the observable inputs (fishing effort) with the observable outputs (yeild or catch). These models have modest data requirements. The input data for these models are a time series of catch and effort from the fishery.

Schaefer's surplus production model and its extensions mainly dominated research in production models for fisheries. Investigations were mainly devoted to (1) model formulation (2) parameter estimation (3) extension to multispecies/ multifleet fisheries and (4) introduction of environmental variables (Schaefer 1954, 1957; Pella and Tomlinson, 1969; Fox, 1970; Schnute, 1977, Pope, 1980; Uhler, 1980; Gulland, 1983; Roff, 1983; Alagaraja, 1984; Tsoa *et al.*, 1985; Ludwig and Walters, 1989; Srinath, 1992; Polachek *et al.*, 1993; Prager, 1994; Laloe, 1995). In this paper the two aspects, namely, the model formulation and parameter estimation are investigated. The choice of appropriate model and estimators of the parameters in the presence of process and observation error are explored. The consequent implications on the management reference points are also examined.

Model formulation and parameter estimation

The surplus production models, also known as biomass dynamic models, assume that the changes in fish population are caused by the interaction of four competing factors, namely, recruitment, growth, reproduction and mortality. The fundamental assumption is that the effects of the three factors, namely, growth, natural mortality and reproduction can be incorporated into a single function and that this is function of stock size only. The change in stock size from one year to the next is thus assumed to be the difference between the biomass dynamic and the catch by the fishery. The simplest production model suggested by Schaefer (*op.cit.*) describes the rate of change in biomass in relation to the biomass and in the absence of fishery it is assumed that

$$DB/dt = rB - (r/K)B^2 \text{ ...... (1)}$$

Where B is the biomass at time t,

r = intrinsic rate of increase in stock

K = maximum population size or the carrying capacity

In the case of exploitation, we have

$$DB/dt = rB - (r/K)B^2 - C \text{ ....(2)}$$

$$= rB - (r/K)B^2 - FB \text{ ....(3)}$$

where, C is the catch and F, the fishing mortality. It is assumed that F is proportional to fishing effort (f) and so that

$F = q.f$ where q is the catchability coefficient

Under equilibrium conditions where $dB/dt = 0$, we have

$$U = C/f = a\text{-}b.f \text{ ......(4)}$$

Where U is the catch per unit effort (CPUE); *a* and *b* are constants.

In practice, this equation is basic to assessment of exploited fish stocks. From this the MSY and f_{MSY} are estimated as $a^2/4b$ and $a/2b$ respectively. This model does not give information of r, K, q and the biomass at different periods of time. The problems associated with the use of equilibrium methods are discussed by various workers (Sissenwine, 1978; Butterworth and Andrew, 1984; Roff and Fairbairn, 1980; Polachek *et.al. op.cit.*). In this paper we deal only with the non-equilibrium version both in the continuous and discrete forms. It is well known that the exploited stocks are affected not only by the fishery dependent factors but also by the fishery independent factors. Fisheries data are often noisy because of the effect of biophysical factors on the stocks and sampling errors in the observation of catch or CPUE. The equilibrium form is not dynamic and does not allow for the environmental variations and thus assumed to be free from noise. It is also known that equilibrium conditions are rarely met in practice. Thus we have considered only the non-equilibrium case. To take into account the uncertainties, stochastic errors are introduced in the process and observation equations. Estimators of the parameters are obtained on the basis of either process error or the observation error. Process error estimators assume that the catch (or the CPUE, the index of abundance) is measured without error and that all the error occurs only in the process with which biomass is generated. Observation error estimators on the other hand assume that there is no stochasticity in population dynamics of fish stocks and the errors are only due to observation of catch or CPUE.

Polachek *et. al., (op.cit.)* have investigated different approaches in estimating the parameters by taking the error components separately. They have not investigated the behaviour of the estimators in the presence of both types of errors. In this paper such an attempt is made and the parameters are estimated using the process error method only. This would also bring out behaviour of the estimators using only one approach, namely, the process error method, when the errors due to observation are also present in the system.

In the presence of stochastic errors both in the process and observation, the Schaefer's model in the discrete and continuous form will be as follows.

In the discrete form we have,

$B_{t+1} - B_t = rB_t - (r/K)B_t^2 - C_t + \varepsilon_t$ ....(5)

$Ct = qftBt + \Phi t$ ....(6)

Here it is assumed that ε_t and Φ_t are independently distributed random variables with zero means and constant variances σ_1^2 and $\sigma_2^{2.}$

Now, we get,

$(B_{t+1} - B_t)/B_t = r - (r/K)\ B_t - qf_t + \varepsilon_t$ ....(7)

(This is obtained by dividing equation 5 by B_t after replacing C_f by equation 6 and taking $\varepsilon_t = (\varepsilon_t - \Phi_t)/B_t$.

Since B_t, the stock biomass is not known the observable proxy for it namely, U_t the catch per unit effort can be used to build the equation.

$U_t = qf_t + \Phi'_t$ ....(8)

Where $\Phi'_t = \Phi_t/f_t$ substituting this in the above equation and after rearranging we get

$(U_{t+1}/U_t) = a_1 + a_2U_t + a_3f_t + \theta_t$ ....(9) $\rightarrow$ model A

where $a_1 = (1 + r)$; $a_2 = -(r/Kq)$ and $a_3 = -q$

and $\theta_t = f(U_t, \varepsilon_t, \Phi_t)$

In the continuous form:

$dB_t/dt = rB_t - (r/K)Bt^2 - C_t + \delta_t$ ....(10)

$C_t = qf_tB_t + \eta t$ ....(11)

$U_t = C_t/f_t = qB_t + \eta'_t$ ....(12)

Substituting C_t in the first equation and integrating over (t, t+1) we get,

$\text{Log}\ (B_{t+1}/B_t) = r-(r/K)B_t - qf_t + \zeta't$ ....(13)

Where B_t and f_t are the average biomass and effort in the interval (t, t+1)

$$\text{where } \zeta't = \int_{t}^{t+1} (\delta_t - \eta_t)\ B_t d_t$$

Since B_t is unobservable we can express the above equation in terms of U_t as follows:

$\text{Log}\ (U_{t+1}/U_t) = b_1 + b_2U_t + b_3f_t + p_t$ ....(14)

Where $p_t = f(U_t, f_t, \zeta t)$ and $b_1 = r$, $b_2 = -r/K$; $b_3 = -q$

It may be noted here that MSY = r.K/4 and $f_{msy} = 0.5.r.q$ and U_t and f_t are the time averages and can be considered as the catch per unit effort and the effort during the year. Since U_t and U_{t+1} are the instantaneous rates they are not usually observed and hence following the approximation of Schnute *(op.cit.)* where $U_t = \sqrt{(U_tU_{t+1})}$ and we have

$\text{Log}(U_{t+1}/U_t) = b_1 + b_2(U_{t+1} + U_t)/2 + b_3(f_{t+1} + f_t)/2 + p_t$ .... (15) $\rightarrow$ model B

Where $p_t = (p_{t+1} + p_t)/2$

Analytically it may not be possible to compare these two models and hence a Monte Carlo bootstrap simulation was made by first generating a population series and a catch series with different error levels for the stock and the catch. The population was generated using the discrete time version of the Schaefer's model with the following inputs r = 0.45; K = 1500; q = 0.000254 for error levels σ_1 = 0, 25, 50, 75, 100 and σ_2 = 0, 10, 20, 30, 40 with initial value of the stock biomass B_0 = 1000.

To generate the catch, the effort data was taken from Miyabe (1989). The data was generated for 36 values of effort. The data was simulated for all the 25 combinations of the error terms. The reference data set is the one with 0 level of error term. The bootstrap regression as suggested by Wu (1986) was used to estimate the parameters and the biases. The number of bootstraps is 1000. The bootstrap estimates the relative bias in r, K,q, MSY and fMSY and the estimated coefficient of variation in r, K and q are given in tables. The bootstrap relative bias in this case is given by $(x - x_0) * 100/x_0$ where x is the average bootstrap estimate (mean over number of bootstrap samples) and x+0+ is the value of the parameter. The coefficient of variation is computed as bootstrap standard deviation of the estimate divided by the bootstrap estimate of the parameter multiplied by 100.

Results and discussion

The bootstrap estimates of relative bias in the estimates of parameters r, K, q (prefixed as 'bias') along with the coefficient of variation (prefixed as 'cv') and those of the estimates of MSY and fMSY are presented in Tables 1 to 6.

It can be seen from Table 1 that when both process and observation errors are zero, model A estimates with almost zero bias but not the model B. The reason for this could be attributed to the fact that the population simulated was based on the discrete form of the Schaefer model and not on the basis of the continuous form. However, the bias and the relative variation in the estimates of the parameters are not large enough and so it is assumed that further comparison would bring out differences in the models. From the tables of the relative bias in the esti-

TABLE 1. *Relative bias in estimates when error in catch equations is 0 (The first row heading is the error levels in the stock equation)*

	0		25		50		75		100	
	A	B	A	B	A	B	A	B	A	B
bias r	0	2.79	42.86	28.68	3.21	-29.35	70.88	-4.54	9.88	-46.99
bias k	0	-12.00	-16.13	-6.20	88.40	175.00	-51.87	46.60	0.07	184.07
bias q	0	4.84	18.36	19.24	-17.56	-7.36	140.20	66.40	18.24	-9.32
cv r	0	3.13	12.54	22.82	30.20	51.25	22.80	43.16	25.38	46.09
cv k	0	3.11	13.75	63.04	141.00	305.23	25.48	390.45	44.90	264.75
cv q	0	2.92	15.04	22.67	46.57	46.76	21.97	39.79	30.51	48.26

TABLE 2. *Relative bias in estimates when error in catch equation is 10 (The first row heading is the error levels in the stock equation)*

	0		25		50		75		100	
	A	B	A	B	A	B	A	B	A	B
bias r	182.57	26.06	430.57	307.88	141.04	-16.99	203.06	119.70	118.57	32.83
bias k	-56.86	258.66	-78.13	-13.80	-4.70	354.00	-55.00	-24.40	115.20	71.40
bias q	155.50	91.04	441.56	332.04	110.16	-8.40	289.32	249.44	63.84	114.36
cv r	13.03	52.68	28.94	47.24	17.63	61.74	46.79	70.40	25.55	50.82
cv k	34.21	1314.7	41.77	538.75	45.53	634.50	125.15	272.81	505.50	323.38
cv q	29.85	47.05	36.19	49.11	34.92	66.59	54.71	67.20	65.38	51.98

TABLE 3. *Relative bias in estimates when error in catch equation is 20 (The first row heading is the error levels in the stock equation)*

	0		25		50		75		100	
	A	B	A	B	A	B	A	B	A	B
bias r	536.91	169.46	252.74	5.04	243.51	39.72	554.28	106.20	332.86	123.04
bias k	-78.47	44.07	-71.13	443.00	-71.53	57.93	-48.13	74.47	-76.47	-43.47
bias q	699.72	392.80	330.16	55.00	319.84	159.48	828.84	390.32	636.40	375.92
cv r	33.14	53.39	17.16	63.16	16.15	40.86	51.18	81.59	30.91	39.31
cv k	120.12	1033.0	35.88	888.95	40.28	285.18	406.40	657.24	154.39	275.00
cv q	49.44	54.07	30.36	72.31	30.76	58.91	69.70	78.01	40.84	47.11

TABLE 4. *Relative bias in estimates when error in catch equation is 30 (The first row heading is the error levels in the stock equation)*

	0		25		50		75		100	
	A	B	A	B	A	B	A	B	A	B
bias r	171.67	5.46	1262.00	131.83	1214.10	212.95	2997.00	448.00	388.38	71.9
bias k	1.07	370.07	-88.07	187.80	-71.67	-48.20	-76.20	116.70	-53.07	193.4
bias q	154.00	48.68	2185.10	564.56	2352.70	674.60	6387.00	1063.00	372.20	211.5
cv r	24.88	71.02	53.49	62.93	56.97	55.29	69.72	61.79	30.16	62.0
cv k	221.83	490.28	154.75	371.34	698.30	187.51	737.53	913.57	393.32	577.6
cv q	64.17	83.56	64.29	64.71	66.94	60.11	77.17	64.22	53.78	68.2

TABLE 5. *Relative bias in estimates when error in catch equation is 40 (The first row heading is the error levels in the stock equation)*

	0		25		50		75		100	
	A	B	A	B	A	B	A	B	A	B
bias r	668.77	200.15	431.80	81.43	1100.00	175.89	903.60	176.58	593.10	165.0
bias k	-86.50	-28.60	-55.60	61.30	-68.40	55.60	-90.60	-28.20	-23.20	-10.6
bias q	1490.70	570.52	617.44	221.56	2440.00	629.40	228.30	599.00	1125.00	429.3
cv r	43.08	60.97	33.98	53.10	74.13	49.90	32.46	66.31	48.65	48.8
cv k	126.60	349.57	678.55	387.70	380.55	572.45	57.44	259.97	1451.40	381.8
cv q	45.18	58.93	49.03	79.19	81.55	69.14	43.45	75.21	65.10	63.2

mates of r, K, q, MSY and f_{MSY} the following observations could be made.

The two models react differently in the presence of both process and observation errors.

The coefficient of variation of the estimates of the parameters obtained from model A was lower than those from model B.

At higher levels of the process errors the model B tends to estimate q with relatively lesser bias than model A.

Both the models overestimated r and q in most of the cases and in general the relative bias in these estimates obtained from model B were lower compared to that obtained from model A.

The maximum sustainable yield tended to be overestimated from both the models and the models under estimated the optimal effort.

Thus from the point of view of estimating the basic parameters of the production model, namely, r, K and q, the continuous form (model B) of the

TABLE 6. *Table showing the relative bias in MSY and* f_{MSY}

	0		25		50		75		100	
	A	B	A	B	A	B	A	B	A	B
Error in catch =	0									
MSY	0	-1.62	19.81	20.71	94.44	94.38	-17.75	39.94	9.96	50.77
f_{MSY}	0	-2.65	20.70	7.92	25.19	-23.74	-28.84	-42.62	-7.07	-41.46
Error in catch =	10									
MSY	21.70	352.15	16.02	251.59	27.74	277.54	34.96	65.95	380.12	127.68
f_{MSY}	10.59	-34.01	3.72	-5.58	14.68	-9.37	-22.16	-37.13	33.65	-38.03
Error in catch =	20									
MSY	37.14	288.20	1.59	471.19	-2.21	120.67	239.55	259.82	1.87	26.09
f_{MSY}	-20.36	-45.32	-18.00	-32.23	-16.99	-46.15	-58.56	-57.74	-41.22	-53.14
Error in catch =	30									
MSY	174.57	395.72	62.53	567.06	272.23	62.08	638.77	1087.3	129.26	404.41
f_{MSY}	6.96	-28.74	-40.40	-65.12	-46.41	-59.60	-52.15	-52.92	3.46	-44.83
Error in catch =	40									
MSY	4.04	114.31	136.12	313.64	3401.2	329.10	-5.28	98.58	431.14	136.79
f_{MSY}	-51.66	-55.07	-25.87	-43.58	-54.69	-62.18	-57.72	-60.43	-43.44	-49.80

Schaefer's model seems to be better choice than the discrete form. Although model A resulted in estimates with lesser coefficient of variation, because of the larger magnitudes of the bias in the estimates precludes the choice of the discrete form.

However, from the management point of view both models tended to over estimate MSY and under estimate f_{MSY}. In this, nevertheless, the discrete form was observed to outperform the continuous form because in general the bias in the estimates of MSY and f_{MSY} were lower for the discrete form. Thus, we have conflicting options before us. Production models tend to estimate some quantities much more precisely than others. For most stocks, the marine biological reference points (MSY, f_{MSY}) are estimated relatively precisely (Prager, 1994). The estimates of stock level and fishing mortality are usually estimated less precisely. This is due to the fact that q (the catchability coefficient) is imprecisely estimated. According to Prager (1994) if a parameterization involving K and r is used in fitting, the estimates of these quantities are usually quite imprecise. However, since they are correlated the corresponding estimates of MSY and the optimum effort can none the less be quite precise. These observations seem to be in good agreement with the results obtained in the case of both the models used. Ludwig (1980) pointed out that if random fluctuations are taken into consideration, the assessment of management strategies was more complicated. While examining the alternative harvesting strategies for three laws of population dynamics, namely, Beverton and Holt, the logistic model and the Pella-Tomlinson model, also found out that the results of the harvesting strategies changes with the noise level in the population and also depended on the type of the model used. The results of Ludwig and Walters (1989) and Polacheck *et.al.* (1993) are not directly comparable with the present findings as they have considered different manifestations of the surplus production models. It is worth noting that the conclusions on the effect of process errors and observation errors are, in general, in close agreement with the earlier similar studies. This study mainly attempted to evaluate the performance of the continuous and discrete forms of the Schaefer's model as given by Schnute (1977) and recommends, in general, the use of the continuous form for estimation of r, K and q. But from the management point of view the discrete form would be the better choice.

In conclusion, it may be mentioned that the ability to choose between different formulations may be driven by the conflicting interests. In this context it may be noted that the models based on simple equations without complete biological interoperation such as the model proposed by Galto and Rinaldi (*op.cit.*), the relative response model of Alagaraja (1984), Roff (1983) and Srinath (1992) may be quite useful in describing the fishery much more accurately and realistically for a given data set.

References

Alagaraja, K. 1984. Simple methods for estimation of parameters for assessing exploited fish stocks. *Indian J. Fish.,* **31**: 177-208

Butterworth, D.S. and P.A. Andrew 1984. Dynamics catch-effort models for the hake stocks in ICSEAF Divisions 1.3 to 2.2, *Colln Scient. Pap. Int. Commn SE Atl. Fish.,* **11**(1): 29-58

Efron, B and R. Tibshirani 1986. Bootstrap methods for standard errors, confidence intervals and other measures of statistical accuracy. *Statistical Science,* 1: (54)77

Fox, W.W. 1970. An exponential surplus yield model for optimizing exploited fish population. *Trans. Amer. Fish Soc.,* **99**: 80-88

Fox, W.W. 1975. Fitting the generalized stock production model by least squares and equilibrium approximations. *Fish Bull.* US, **73**: 23-37

Gulland, J.A. 1983. *Fish stock assessment: a manual of basic methods.* Chichester U.K. Wiley Interscience, FAO/Wiley Series on food and agriculture, **1**: 223 pp.

Laloe, F. 1995. Should surplus productions models be fishery description tools rather than biological models? *Aquat. Living Resour.,* **8**: 1-16

Ludwig, D and C.J. Walters 1989. A robust method for parameter estimation from catch and effort data. *Can. J. Fish. Aquat. Sci.,* **46**: 137-144.

Miyabe, N. 1989. Preliminary stock assessment of pacific big eye tuna. In: *Report of the 3rd Southeast Asian Tuna Conference Bali, Indonesia, 22-24 August 1988.* IPTP/89/GEN/17

Pella, J.J. and P.K. Tomlinson 1969. A generalized stock production model. *Bull. IATTC,* **13**: 419-496.

Polacheck, T., R. Hillborn and A.F. Punt 1993. Fitting surplus production models: comparing methods and measuring uncertainties. *Can. J. Fish. Aquat. Sci.,* **50**: 2597-2607.

Pope, J.G. 1980. Assessment of multispecies resources. In: *Selected lectures from the CIDA/FAO/CECAF Seminar on fishery resource evaluationl. FAO/TF/INT 180(C) Can. Suppl. 166 pp.*

Prager, M.H. 1994. A suite of extensions to a non equilibrium surplus production model. *Fish. Bull.,* **92**: 374-389

Roff, D.A. and D.J. Fairbairn 1980. An evaluation of Gulland's method for fitting the Schaefer's model. *Can.J. Fish. Aquat. Sci.,* **37**: 1229-1235.

Roff, D.A. 1983. Analysis of catch-effort data, comparison of three methods. *Can.J. Fish. Aquat. Sci.,* **40**: 1496-1506

Schaefer, M.B. 1954. Some aspects of the dynamics of population important to the management of the commercial marine fisheries. *Bull. IATTC*-1: 26-56

Schaefer, M.B. 1957. A study of the dynamics of the fishery for Yellowfin tuna in the eastern tropical Pacific Ocean. *Bull IATTC,* **2**: 247-285.

Schnute, J. 1977. Improved estimates from the Schaefer production model. *Can. J. Fish. Aquat. Sci.,* **42**: 583-603

Sissenwine, M.P. 1978. Is MSY an adequate foundation for optimum yield?. *Fisheries,* **3**: 22-42

Srinath, M. 1992. Catch-effort relationship in Pacific bigeye tuna fishery. *Naga,* **15**: 29-30

Tsoa, E., W.E. Schrank and N. Roy 1985. Generalizing fisheries models: an extension of Schaefer's analysis. *Can. J. Fish. Aquat. Sci*., **42**: 44-50

Uhler, R.S. 1980. Least squares regression estimates of the Schaefer production model: Some Monte Carlo simulations. *Can. J. Fish. Aquat. Sci.,* **37**: 1284-1294

Wu, C.F.J. 1986. Jackknife, Bootstrap and other resampling methods in regression analysis. *The Annals of Statistics,* **14** (4): 1261-1295